

M.I.T. LIBRARIES - DEWEY

Digitized by the Internet Archive
in 2011 with funding from
Boston Library Consortium Member Libraries

<http://www.archive.org/details/informationalher00smit>

HB31
M415
no.96-9

II

**working paper
department
of economics**

INFORMATIONAL HERDING AS EXPERIMENTATION DÉJÀ VU

Lones Smith
Peter Sørensen

96-9

Mar. 1996

**massachusetts
institute of
technology**

**50 memorial drive
cambridge, mass. 02139**



INFORMATIONAL HERDING AS EXPERIMENTATION \checkmark VU

**Lones Smith
Peter Sørensen**

96-9

Mar. 1996

MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

JUL 18 1996

LIBRARIES

*Informational Herding as Experimentation Déjà Vu**

Lones Smith[†]

Department of Economics, M.I.T

Peter Sørensen[‡]

Nuffield College and M.I.T.

March 27, 1996

Abstract

We prove that the recently proposed informational herding models are but special cases of a standard single person experimentation model with myopia. We then re-interpret the incorrect herding outcome as a familiar failure of complete learning in an optimal experimentation problem.

We next explore that experimentation model with patient individuals — or equivalently, the observational learning model where individuals care about successors. As such, we find that even when individuals internalize the herding externality in this fashion, incorrect herds and incomplete learning still obtain. We note that this outcome can be implemented as a constrained social optimum when decentralized by transfers.

*The proposed mapping in this paper first appeared in July 1995 as a later section in our companion paper, “Pathological Outcomes of Observational Learning.” For the current full length paper, we thank Abhijit Banerjee, Christopher Wallace, and the MIT theory lunch for comments. Smith gratefully acknowledges financial support for this work from the NSF, and Sørensen equally thanks the Danish Social Sciences Research Council.

[†]e-mail address: lones@lones.mit.edu

[‡]e-mail address: peter.sorensen@nuf.ox.ac.uk

And out of old booke, in good faithe,
Cometh al this new science that men lere.

— Geoffrey Chaucer¹

1. INTRODUCTION

The last few years has seen a flood of research on a paradigm known as informational herding. We ourselves have actively participated in this research herd (Smith and Sørensen (1996a), or simply SS) that was sparked independently by Banerjee (1992) and BHW: Bikhchandani, Hirshleifer, and Welch (1992). The context is seductively simple: An infinite sequence of individuals must decide on an action choice from a finite menu. Everyone has identical preferences and menus, and each may condition his decision both on his (endowed) private signal about the state of the world and on all his predecessors' decisions; however, observation of their private signals is verboten.

If these private signals have bounded power, it is known that a ‘herd’ eventually arises, but is not always correct — namely, after some point, all make the identical choice, possibly unwise. This simple pathological outcome has understandably attracted much fanfare.

In SS we explored, embellished, and fully fleshed out this story, and found that herding is quite robust. BHW’s kindred notion of *cascades* is not as resilient: While beliefs converge upon a limit where only one action is taken with probability one (a *limit cascade*), this need not occur in finite time. The potential for bad herds is also not without caveat: Absent a uniform bound on the strength of the individuals’ private signals, only correct herds arise. Finally, in a world with multiple preference types, a *confounded learning* outcome might occur, where the lesson of history is forever mixed, and private signals always decisive.

A Possible Link, and a Puzzle. Robustness aside, this paper pursues a different train of thought: Is informational herding fundamentally a *new* phenomenon? We have been piqued by its similarity to the familiar failure of complete learning in an optimal experimentation problem. One classic example is Rothschild’s (1974) analysis of the *two-armed bandit*: An infinite-lived impatient monopolist optimally experiments with two possible prices each period, with purchase chances for each price being fixed, unknown draws from $[0, 1]$. Rothschild showed that the monopolist would (i) eventually settle down on one of the prices, and (ii) with positive probability select the less profitable price.

Aiming for more than a casual analogy between this outcome and herding, we must recast the observational learning paradigm as a single person optimization problem. This

¹See The Assembly of Fowles. Line 22.

suggests considering the *forgetful experimenter*, who each period receives a new informative signal, takes an optimal action, and then promptly forgets his signal; the next period, he can reflect only on his action choice. Alas, this is neither a very satisfying model of rationality, nor can it be the basis for an optimal experimentation problem. How then can an experimenter not observe the private signals, and yet take informative actions?

The Second Puzzle, and a Resolution. An interesting sequel to Rothschild's work was McLennan (1984), who permitted the monopolist the flexibility to charge one of a continuum of prices, but assumed only two possible linear purchase chance 'demand curves'. When the demand curves crossed, he found that the monopolist may well settle down on the suboptimal uninformative price.

Rothschild's and McLennan's models give examples of *potentially confounding actions*, as introduced in EK: Easley and Kiefer (1988). In brief, these are actions that are optimal for *unfocused* beliefs for which they are invariants (i.e. taking the action leaves the beliefs unchanged). Of particular significance is the proof in EK (on page 1059) that with finite state and action spaces, there will *generically* not exist any potentially confounding actions, and thus complete learning must arise.² Rothschild and McLennan might be seen as separate anticipations of EK's general insight. Rothschild escapes it by means of a continuous state space, whereas McLennan resorts to a continuous action space. Yet there appears no escape for the herding paradigm, where both flavors of incomplete learning, limit cascades and confounded learning, generically arise with two actions and two states.

Our resolution of these puzzles respects the quintessence of the herding paradigm that predecessors' signals are hidden from view. In a nutshell, we imagine that individuals do not act for themselves, but rather furnish optimal history-contingent 'decision rules' for agent machines that automatically map any realized private signal into an action choice. *With this device, informational herding is properly understood as a new (camouflaged) context for an old phenomenon: incomplete learning by an infinite-lived experimenter.*

Herding with Forward-Looking Behavior. The herding outcome is a striking example of an informational externality: While all individuals collectively know enough to fully determine the state of the world, it is aggregated rather poorly. For in a herd, most individuals almost surely take an action which reveals almost none of their information. Late-comers ideally prefer that their predecessors had better signalled their information with more revealing actions, but early individuals clearly have no incentive to do so.

Notwithstanding the motivation of this research, some may therefore find the second

²For instance, payoff assignments in a one-armed bandit — where the safe arm is a potentially confounding action — are not generic in \mathbb{R}^2 .

half of this paper more compelling, where we investigate the herding externality with forward-looking behavior. In fact, having recast the problem as one of optimal individual experimentation, it is not implausible that learning is incomplete simply because of the myopia. For as is well known, a rational experimenter is willing to forgo some current payoff in order to secure knowledge relevant for future payoffs and decisions. We prove in fact that in the experimentation model without myopia, learning is still not complete.³

Herding itself is best captured by an equivalent *social planner's problem*, where the goal is to maximize the present discounted value of individuals' welfares. The social planner likewise will sacrifice early payoffs for the informational benefit of successors. Such an outcome can be decentralized as a constrained social optimum by means of a simple set of budget balance transfers that encourage experimentation. Yet we find that *the herding externality is only imperfectly mitigated by the social planner: Incorrect herds still obtain, but with chances vanishing as the discount factor converges to one.*

Overview. Section 2 outlines a general observational learning model. In section 3, we re-interpret the herding paradigm as one of optimal single-person experimentation. Section 4 characterizes the limiting beliefs of that experimenter, and then reverts to the social planner's herding problem for the characterization of the limiting actions. A conclusion gives a broader perspective on our findings. Longer but essential proofs are appendicized.

2. OBSERVATIONAL LEARNING MODEL

In this section we set up a very general observational learning model, that subsumes SS, and thus BHW and Banerjee (1992), and also happens to cover Lee (1993). This generality facilitates the ensuing mapping into the experimentation literature.

Information. An infinite sequence of individuals $n = 1, 2, \dots$ takes actions in that exogenous order. There is uncertainty about the payoffs from these actions. The elements of the parameter space (Θ, \mathcal{F}) are referred to as *states of the world*. There is a given common prior belief, the probability measure λ over Θ .

Individual n receives a private random signal, $\sigma_n \in \Sigma$, about the state of the world. As demonstrated in SS (Lemma 1), we may assume WLOG that the private signal received by an individual is actually his *private belief*, i.e. we let σ be the measure over Θ which results from Bayesian updating given the prior λ and observation of the private signal. Signals thus belong to Σ , the space of probability measures over (Θ, \mathcal{F}) , and \mathcal{G} is the associated sigma-algebra. Conditional on the state, the signals are assumed to be i.i.d.

³Just as in SS, when private signals are not uniformly bounded in power, learning is complete.

across individuals, distributed according to the probability measure μ^θ in state $\theta \in \Theta$. To avoid trivialities, assume that not all μ^θ are identical, so that some signals are informative. Each distribution may contain atoms, but to ensure that no signal will *perfectly* reveal the state of the world, we insist that all μ^θ be mutually absolutely continuous (a.c.), for $\theta \in \Theta$.⁴

Bayesian Decision-Making. Everyone chooses from a fixed action set A , equipped with the sigma-algebra \mathcal{A} . Action a earns a nonstochastic payoff $u(a, \theta)$ in state $\theta \in \Theta$, the same for all individuals, where $u : A \times \Theta \mapsto \mathbb{R}$ is measurable. It is common knowledge that everyone is rational, i.e. seeks to maximize their expected payoff. Before deciding upon an action, everyone first observes his private signal/belief and the entire action history h .

An individual's Bayes-optimal decision rule uses the observed action history and his own private belief. As in SS, we simply assume that an individual can compute the decision rules of all predecessors, and can use the common prior to calculate the *ex ante* (time-0) probability distribution over action profiles h in either state. Knowing these probabilities, Bayes' rule implies a unique *public belief* $\pi = \pi(h) \in \Sigma$ for any history h .⁵ A final application of Bayes' rule also given the private signal σ yields the *posterior belief* $\rho \in \Sigma$.

Given the posterior belief $\rho \in \Sigma$, individual n picks the action $a \in A$ which maximizes his expected payoff $\bar{u}_a(\rho) = \int_{\Theta} u(a, \theta) d\rho(\theta)$. We assume that an optimal action $a = a(\rho)$ always exists.⁶ The solution defines an *optimal decision rule* $x : \Sigma \mapsto \Delta(A)$, the space of probability measures over (A, \mathcal{A}) . That is, x is an element of the space X of maps from Σ to $\Delta(A)$. A rule x produces an implied distribution $\nu = x(\sigma)$ over actions simultaneously for all private beliefs σ . The optimal x clearly depends on π .

Observational Learning as a Stochastic Process. Since the probability measure of signals σ depends on the state θ , the implied distribution over actions a depends on both θ and the optimal decision rule x . In fact, the density is $\psi(a|\theta, x) \equiv \int x(\sigma)(a) \mu^\theta(d\sigma)$, and unconditional on the state, it is $\psi(a|\pi, x) \equiv \int_{\Theta} \psi(a|\theta, x) \pi(d\theta)$. This in turn yields a distribution over next period public beliefs π_{n+1} . Thus, $\langle \pi_n \rangle$ follows a discrete-time Markov process with state-dependent transition probabilities. The next result is standard:

Lemma 1 *The belief process $\langle \pi_n \rangle$ is a martingale unconditional on the state, converging a.s. to some limiting random variable π_∞ . The limit π_∞ places no weight on point masses on the wrong states of the world.*

⁴See Rudin (1987). Measure μ^1 is a.c. w.r.t. μ^2 if $\mu^2(S) = 0 \Rightarrow \mu^1(S) = 0 \forall S \in \mathcal{S}$, the sigma-algebra on Σ . By the Radon-Nikodym Theorem, a unique $g \in L^1(\mu^2)$ exists with $\mu^1(S) = \int_S g d\mu^2$ for every $S \in \mathcal{S}$. With μ^H, μ^L mutually a.c., 'almost sure' assertions are well-defined without specifying the measure.

⁵If one wishes to pursue this angle, this is the unique Bayesian equilibrium of the 'game'.

⁶Absent a unique solution, we must take an arbitrary measurable selection from the solution correspondence.

OBSERVATIONAL LEARNING MODEL	IMPATIENT EXPERIMENTER MODEL
States $\theta \in \Theta$	States $\theta \in \Theta$
Belief after n individuals π_n	Belief after n observations π_n
Optimal decision rule $x \in X$	Optimal action $x \in X$
Private signal of individual n , σ_n	Randomness in the n th experiment
Action taken by individual, $a \in A$	Observable signal $a \in A$
Density over actions $\psi(a \theta, x)$	Density over observables $\psi(a \theta, x)$
Payoffs (private information)	Payoffs (unobserved)

Table 1: **Embedding.** This table displays how our single-type observational learning model fits into the impatient single person experimentation model.

3. MODEL COMPARISON

A first step in recasting our general model of observational learning as a single person experimentation problem is to respect the individuals' selfishness. Thus, we must study an impatient experimenter with discount factor 0 (no 'active experimentation').

But to avoid a forgetful experimenter, we must regard the observational learning story from a new perspective. Consider individual n , who uses both the public belief π_n and his private signal σ_n in forming and acting upon his posterior beliefs ρ_n . We may separate these two steps using the conditional independence of π_n and σ_n . Mr. n can be regarded as: (i) observing π_n , but *not* his private signal; (ii) optimally determining the rule $x \in X$, and submitting it to an agent 'choice' machine; and (iii) letting that machine observe his private signal and take his action $a \in A$ for him. The ultimate payoff $u(a, \theta)$ is unobserved, lest that provide an additional signal of the state of the world. If private beliefs σ have distribution μ^θ in state θ , the impatient experimenter will choose the same optimal decision rule x described in section 2, resulting in action $a \in A$ with chance $\psi(a|\theta, x)$.

We now precisely describe the single-person experimentation model. The state space is Θ . In period n , the experimenter \mathcal{EX} chooses an action (the rule) $x \in X$. Given this choice, a random observable statistic $a \in A$ is realized with chance $\psi(a|\theta, x)$ in state θ . Finally, \mathcal{EX} updates beliefs using this information alone.⁷ Table 1 summarizes the embedding.

Notice how this addresses both lead puzzles. First, the experimenter never knows

⁷This experimentation model does not strictly fit into the EK mold, where the instantaneous reward in period n depends only on the action and the observed signal, but (unlike here) not on the parameter θ in Θ . This is the structure of Aghion, Bolton, Harris, and Jullien (1991) (ABHJ), where payoffs are not necessarily observed. If we wish, we may simply posit that the experimenter has fair insurance, and simply earns his expected payoff each period rather than his random realized payoff. Then his behaviour will be exactly the same, yet he will not learn anything from the payoff.

the private beliefs σ , and thus does not forget them. Second, the pathological learning outcomes are entirely consistent with EK's generic complete learning result for models with finite action and state spaces. Simply put, *actions do not map to actions but to signals when one rewrites the observational learning model as an experimentation model*. The true action space for \mathcal{EX} is the infinite space X of decision rules.

SS considered two major modifications of the informational herding paradigm. One was to add i.i.d. noise to the individual decision problem. Noise is easily incorporated here by adding an exogenous chance of a noisy signal (random action). SS also allowed for T different types of preferences, with individuals randomly drawn from one or the other type population. Multiple types can be addressed here by simply imagining that \mathcal{EX} chooses a T -vector of optimal decision rules from X^T with (only) the choice machine observing the task and private belief, and choosing the action a as before.

4. PATIENCE

4.1 Special Assumptions

While we have rather generally expressed observational learning as a type of myopic experimentation, our analytical results will require a more restricted model typical of the herding paradigm. For instance, if (Θ, \mathcal{F}) is a continuum subset \mathbb{R} equipped with the Borel sigma-algebra, then the space of mappings X is rather unwieldy: the space of measurable mappings from the measures Σ over Θ to the simplex $\Delta(A)$.

So just as in SS, we assume that $\Theta = \{H, L\}$ (or is more generally finite). High and low states are equilike ex ante, and so have prior $\lambda(L) = \lambda(H) = 1/2$. Private belief σ expresses the chance of state H , and thus Σ is the interval $[0, 1]$, and \mathcal{G} the Borel sigma-algebra over $[0, 1]$. Let $\text{supp}(\mu)$ be the (common) support⁸ of each probability measure μ^θ . If $\text{supp}(\mu) \subseteq (0, 1)$, then we call the private beliefs *bounded*, while if $\text{co}(\text{supp}(\mu)) = [0, 1]$, they are *unbounded*. In that case, arbitrarily strong private signals/beliefs can occur.

Lee (1993) showed that with these restrictions, a continuous action space can easily allow for simple statistical learning: Knowing individual n 's public belief, observation of his action then perfectly reveals his private signal. We thus make the standard lumpiness assumption that $A = \{a_1, \dots, a_M\}$ is finite. We assume that no action is dominated by the others, and this implies the simple interval structure that action a_m is optimal exactly when the posterior ρ is in some sub-interval of $[0, 1]$. We order the actions such that a_m is myopically optimal for posteriors $\rho \in [\bar{r}_{m-1}, \bar{r}_m]$, where $0 = \bar{r}_0 < \bar{r}_1 < \dots < \bar{r}_M = 1$.

⁸As usual, the support of a probability on the Borel-algebra is the smallest closed set of measure 1.

4.2 The Patient Experimenter

Since our experimentation model admits any discount factor $\delta \in [0, 1]$, one naturally wonders what happens with a patient experimenter (when $\delta > 0$). SS shows that the myopic experimenter maps higher signals into higher actions: There are thresholds $0 = \bar{x}_0 \leq \bar{x}_1 \leq \dots \leq \bar{x}_M = 1$ depending on π alone, such that action a_m is strictly optimal when $\sigma \in (\bar{x}_{m-1}, \bar{x}_m)$, and indifference between a_m and a_{m+1} prevails at the knife-edge $\sigma = \bar{x}_m$. This is also true with patience, as (appendicized) Lemma A.1 proves. Intuitively, not only does the interval structure maximize immediate payoffs, but it also ensures the greatest future value of information, by producing the riskiest posterior belief distribution.

We now turn to \mathcal{EX} 's long run behavior. Lemma 1 tells us that beliefs must settle down and that \mathcal{EX} is never dead wrong about the state. The next result is an expression of EK's Theorem 5 that the limiting belief π_∞ precludes further learning.

Proposition 1 (Absorbing Basins) *For each $a_m \in A$, a possibly empty interval exists $J_m(\delta) \subset [0, 1]$, such that when $\pi \in J_m(\delta)$, \mathcal{EX} optimally chooses x which a.s. induces a_m .*

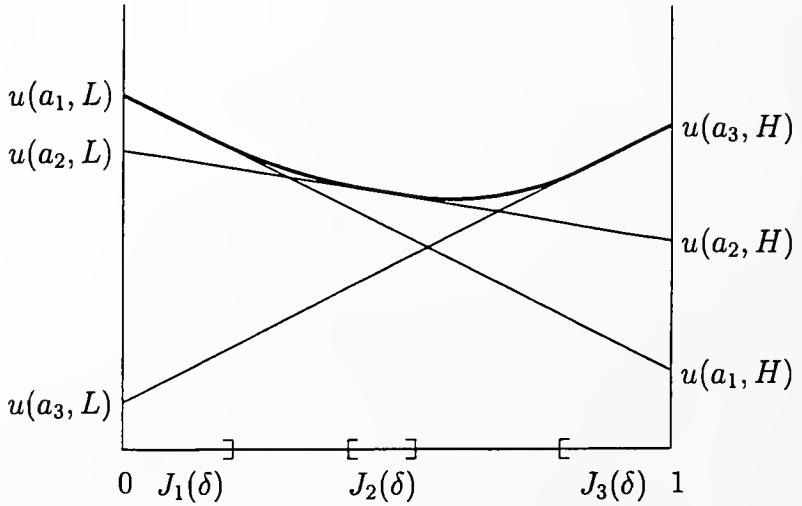
- For all $\delta \in [0, 1]$, the limit belief π_∞ is concentrated on the basins $J_1(\delta) \cup \dots \cup J_M(\delta)$.
- With unbounded private beliefs, $J_1(\delta) = \{0\}$ and $J_M(\delta) = \{1\}$; all other $J_m(\delta)$ are empty.
- If the private beliefs are bounded, then $J_1(\delta) = [0, \underline{\pi}(\delta)]$ and $J_M(\delta) = [\bar{\pi}(\delta), 1]$, where $0 < \underline{\pi}(\delta) < \bar{\pi}(\delta) < 1$. The larger is δ , the smaller are all basins. For large enough δ , all basins disappear except for J_1 and J_M , while $\lim_{\delta \rightarrow 1} J_1(\delta) = \{0\}$ and $\lim_{\delta \rightarrow 1} J_M(\delta) = \{1\}$.

Proof: This characterization of the stationary points of the stochastic process of beliefs $\langle \pi_n \rangle$ directly generalizes the analysis for $\delta = 0$ in SS. See figure 1 for an illustration of how the basins are determined from the shape of the optimal value function. All but the initial limit belief result are established in the appendix. To see why that one is true — that a *limit cascade* must occur, as SS call it — observe that for any belief $\hat{\pi}$ not in any basin, at least two signals in A are realized with positive probability. By the interval structure, the highest such signal is more likely in state H , and the lowest more likely in state L . It follows that next period's belief differs from $\hat{\pi}$ with positive probability. Intuitively, or by the characterization result for Markov-martingale processes in appendix B of SS, $\hat{\pi}$ cannot lie in the support of π_∞ . \square

Proposition 2 (Convergence of Beliefs, or Long Run Learning)

- For unbounded private beliefs, π_∞ is concentrated on the truth for any $\delta \in [0, 1]$.
- With bounded private beliefs, learning is incomplete for any $\delta \in [0, 1]$: Unless $\pi_0 \in J_M(\delta)$, there is positive probability in state H that π_∞ is not in $J_M(\delta)$.
- The chance of incomplete learning with bounded private beliefs vanishes as $\delta \uparrow 1$.

Figure 1: **Typical value function.** Stylized graph of $v(\pi, \delta)$, $\delta > 0$, with three actions.



Proof: Given the absorbing basin characterization of Proposition 1, the unbounded beliefs result follows from Lemma 1 (since π_∞ places no weight on the point belief $\pi = 0$). The bounded beliefs incomplete learning conclusion follows just as in Theorem 1 of SS. We now extend that proof to establish the limiting result for $\delta \uparrow 1$. First, Proposition 1 assures us that for δ close enough to 1, π_∞ places all weight in $J_1(\delta)$ and $J_M(\delta)$. The likelihood ratio $\ell_n \equiv (1 - \pi_n)/\pi_n$ is a martingale conditional on state H . Because all private beliefs σ have likelihood ratio $(1 - \sigma)/\sigma$ bounded above by some $\bar{\ell} < \infty$, the sequence $\langle \ell_n \rangle$ is bounded above by $\bar{\ell}(1 - \underline{\pi}(\delta))/\underline{\pi}(\delta)$, and the mean of ℓ_∞ must equal its prior mean $(1 - \pi_0)/\pi_0$. Since $\lim_{\delta \rightarrow 1} \underline{\pi}(\delta) = 0$, the weight that π_∞ places on $J_1(\delta)$ in state H must vanish as $\delta \rightarrow 1$. \square

Observe how incomplete learning to some extent plagues even an extremely patient \mathcal{EX} . So this problem does not fall under the rubric of EK's Theorem 9, where it is shown that if the optimal value function v is *strictly convex* in beliefs π , learning is complete for δ close enough to 1. For here, \mathcal{EX} optimally behaves myopically for very extreme beliefs: $v(\pi) = \bar{u}_{a_1}(\pi)$ for π near 0, and $v(\pi) = \bar{u}_{a_M}(\pi)$ for π near 1, both *affine* functions. This points to the source of the incomplete learning: lumpy signals rather than impatience.

4.3 Forward-Looking Informational Herding

Let us return to the view of this as the informational herding paradigm. Suppose first that everyone is altruistic, but subject to the standard herding informational constraints. If the realized payoff sequence is $\langle u_k \rangle$, suppose that every individual $n = 1, 2, \dots$ seeks to maximize $E[(1 - \delta) \sum_{k=n}^{\infty} \delta^{k-n} u_k | \pi]$, where we define $E[f | \pi] \equiv \sum_{\theta \in \Theta} \pi(\theta) f(\theta)$. Then the

solution to \mathcal{EX} 's problem yields a perfect Bayesian equilibrium of this model.

The Planner's Problem. Now let us add a shade of economic realism, and posit selfish behavior. Reinterpret the patient experimenter's problem as that of an informationally constrained social planner \mathcal{SP} trying to maximize $(1 - \delta) \sum_{n=1}^{\infty} \delta^{n-1} u_n$, the expected discounted average welfare of the individuals in the herding model. Observe how this perfectly aligns the \mathcal{SP} 's and \mathcal{EX} 's objectives. To respect that key herding informational assumption that actions are observed and signals unobserved, assume further that the \mathcal{SP} neither knows the state nor can observe the individuals' private signals, but can both observe and punish/reward any actions taken.

We are now positioned to reformulate the learning results of the last section at the level of actions. The *Overturning Principle* of SS generalizes to this case: Lemma A.2 establishes that for π near $J_m(\delta)$, actions other than a_m will push the updated public belief far from its current value. This at once yields the following corollary to Proposition 2.

Proposition 3 (Convergence of Actions, or Herds)

- For unbounded private beliefs, a herd eventually starts on the correct action.
- With bounded private beliefs, a herd on some action eventually starts. Unless $\pi_0 \in J_M(\delta)$, a herd arises on an action other than a_M with positive chance in state H for any $\delta \in [0, 1)$.
- The chance of an incorrect herd with bounded private beliefs vanishes as $\delta \uparrow 1$.

It is no surprise that \mathcal{SP} ends up with full learning with unbounded beliefs, for even selfish individuals will. More interesting is that \mathcal{SP} optimally incurs the risk of an ever-lasting incorrect herd. Herding is truly a robust property of the observational learning paradigm.

Optimal Transfers. How does \mathcal{SP} steer the choices away from the myopic solution to \mathcal{EX} 's problem? Let \mathcal{SP} tax or subsidize the actions according to the following simple scheme. Given the current public belief π , if the individual takes action a he receives the (possibly negative) transfer $\tau(a|\pi)$. Faced with such incentives, our proof in Lemma A.1 that individuals optimally choose private belief threshold rules is still valid.

Given public belief π , and optimal rule x^* for \mathcal{SP} , how are the transfers set? The private belief σ is mapped into the posterior $\rho(\pi, \sigma) = \pi\sigma / [\pi\sigma + (1 - \pi)(1 - \sigma)]$. The selfish herder's threshold \bar{x}_m must solve the indifference equation $\bar{u}_{a_m}(\rho(\pi, \bar{x}_m)) + \tau(a_m|\pi) = \bar{u}_{a_{m+1}}(\rho(\pi, \bar{x}_m)) + \tau(a_{m+1}|\pi)$. So, the difference $\tau(a_{m+1}|\pi) - \tau(a_m|\pi)$ alone matters for how individuals trade-off between the two actions, and the \mathcal{SP} can ensure that the threshold belief is optimally chosen, $\bar{x}_m = \bar{x}_m^*$, by suitably adjusting this difference.

As incentives are unchanged if a constant is added to all M transfers, we can also insist that \mathcal{SP} achieve *expected budget balance* each period: i.e. the expected contribution from everyone is zero, or $0 = \sum_{m=1}^M \psi(a_m|\pi, \bar{x})\tau(a_m|\pi)$. This uniquely determines the transfers.

We would dearly like to provide a crisp characterization of these transfers — but after much work, we believe that this is impossible. Clearly, \mathcal{SP} will not tax or subsidize actions if his desired one will be chosen anyway, i.e. for $\pi \in J_i(\delta) \subseteq J_i$. Conversely, if \mathcal{SP} wishes that actions be chosen with discretion as when $\pi \notin J_i(\delta)$, then some transfers intuitively will differ from zero, because myopic and patient behavior do not coincide. Rigorously, as $\pi \in J_i(\delta) \subset J_i$ is a strict inclusion for high enough discount factors by Claim A.7, transfers are not identically zero for a sufficiently patient \mathcal{SP} . Beyond that, we can only say that ‘experimentation’ (making non-myopic choices) is rewarded, as it provides information to successors. What these ‘high-information’ choices are is rather hard to tell.

5. CONCLUSION

This paper has discovered and explored the fact that informational herding is simply incomplete learning by a single experimenter, suitably concealed. Yet herding has clearly achieved a measure of ‘market penetration’ that has eluded the experimentation literature. Aside from its simple economics, we believe this owes to its focus on the zero discounting case, thus avoiding the heavy dynamic optimization machinery.

Our mapping, recasting everything in rule space, has led us to an equivalent social planner’s problem, an oft-performed exercise for economic models. The *revelation principle* in mechanism design also employs a ‘rule machine’. But in multi-period, multi-person models with uncertainty, the planner must respect the agents’ belief filtrations. Our rule machine approach succeeds by exploiting the martingale property of public beliefs.

This property extends beyond action observation models to those where only an imperfect signal of one’s posterior belief is observed — again, provided the entire history of such signals is observed. But once we relax the assumption that the history is perfectly observed, as in Smith and Sørensen (1996b), the public belief process (howsoever defined) ceases to be a martingale — a fact that we are exploring in a work in progress. As a result, a mapping of these models into the experimentation literature no longer appears possible.

Finally, we and others have correctly attributed bad herds to the lumpy transmission of social information (a finite action space). As in EK, incomplete learning at the very least requires an optimal action x for which unfocused beliefs are invariant, i.e. the distribution $\psi(a|\theta, x)$ of signals a is the same for all states. Such an invariance is clearly easier to satisfy with fewer available signals, and not surprisingly herding and all published experimentation pathologies that we have seen assume a finite (vs. continuous) signal space. For instance, Rothschild (1974), McLennan (1984), and ABHJ’s example are all binary signal models.

A. APPENDIX

A.1 Preliminary

A strategy s_n for period n is a map from Σ to X . It prescribes the rule $x_n \in X$ which must be used, given belief π_n . The experimenter chooses a strategy profile $s = (s_1, s_2, \dots)$, which in turn determines the stochastic evolution of the model — i.e. a distribution over the sequences of realized actions, signals, payoffs, and beliefs.

Our analysis is inspired by ABHJ and sections 9.1–2 of Stokey and Lucas (1989). The value function $v(\cdot, \delta) : \Sigma \mapsto \mathbb{R}$ for the experimentation problem with discount factor δ is $v(\pi, \delta) = \sup_s E[(1 - \delta) \sum_{n=1}^{\infty} \delta^{t-1} u_n | \pi]$, where the expectation uses the distribution of processes implied by s . Recall that $\bar{u}_m(\pi) = \pi u(a_m, H) + (1 - \pi) u(a_m, L)$ denotes the expected payoff from a_m at belief π . Since \bar{u}_m is affine, we have the Bellman operator T_δ

$$T_\delta v(\pi) = \sup_{x \in X} \left\{ \sum_{a_m \in A} \psi(a_m | \pi, x) [(1 - \delta) \bar{u}_m(q(\pi, x, a_m)) + \delta v(q(\pi, x, a_m))] \right\} \quad (\text{A-1})$$

where $q(\pi, x, a)$ is the Bayes-updated belief from π when a is observed and rule x is applied.

Note that for $v \geq v'$ we have $T_\delta v \geq T_\delta v'$. As is standard, T_δ is a contraction, and $v(\cdot, \delta)$ is its unique fixed point in the space of bounded, continuous, weakly convex functions.

Lemma A.1 (Interval Structure) *For any π , one optimal rule $x \in X$ is described by thresholds $0 = \bar{x}_0 \leq \bar{x}_1 \leq \dots \leq \bar{x}_M = 1$ such that action a_m is taken when $\sigma \in (\bar{x}_{m-1}, \bar{x}_m)$, and when $\sigma = \bar{x}_m$ there is randomization between a_m and a_{m+1} .*

Proof: We prove that any rule x violating this interval structure can be strictly improved upon. For $m_1 < m_2$, let Σ_i ($i = 1, 2$) be those signals in Σ which are mapped with positive probability into a_{m_i} . Assume to the contrary that the sets are *not* ordered as $\Sigma_1 \leq \Sigma_2$.

If $q(\pi, x, a_{m_1}) > q(\pi, x, a_{m_2})$, then given our ordering of actions, payoffs are improved (with no loss of information) by reinapping signals leading to a_{m_1} into a_{m_2} , and vice versa.

Next, assume that $q(\pi, x, a_{m_1}) \leq q(\pi, x, a_{m_2})$. For any $\bar{x} \in (0, 1)$, define $\tilde{\Sigma}_1(\bar{x}) \equiv (\Sigma_1 \cup \Sigma_2) \cap [0, \bar{x}]$ and $\tilde{\Sigma}_2(\bar{x}) \equiv (\Sigma_1 \cup \Sigma_2) \cap [\bar{x}, 1]$. Consider then the modified rule \tilde{x} which equals x , except that a_{m_i} is taken for signals in $\Sigma_i(\bar{x})$, and where \bar{x} satisfies $\psi(a_{m_i} | \pi, x) = \psi(a_{m_i} | \pi, \tilde{x})$ (it may be necessary for \tilde{x} to randomize over the two actions at signal \bar{x} to accomplish that). Since signals more in favor of state L are mapped into a_{m_1} under \tilde{x} , we find $q(\pi, \tilde{x}, a_{m_1}) \leq q(\pi, x, a_{m_1})$, and similarly $q(\pi, \tilde{x}, a_{m_2}) \geq q(\pi, x, a_{m_2})$. Thus, \tilde{x} implies a mean preserving spread of the updated belief versus x , and since the value function is weakly convex, this weakly improves the value in the Bellman equation $T_\delta(v) = v$. \square

A.2 Proof of Proposition 1

The proposition is established in a series of claims.

Claim A.1 (Basins) *For each $a_m \in A$, a possibly empty interval $J_m(\delta)$ exists, such that when $\pi \in J_m(\delta)$, \mathcal{EX} optimally chooses x with $\text{supp}(\mu) \subseteq [\bar{x}_{m-1}, \bar{x}_m]$, i.e. a_m occurs a.s. (learning stops). For any $\delta \in [0, 1)$, $0 \in J_1(\delta)$ and $1 \in J_M(\delta)$.*

Proof: For the first half, we really need only prove that $J_m(\delta)$ must be an interval. If $\pi \in J_m(\delta)$, then a_m is the optimal choice, and the value is $v(\pi, \delta) = v_0(\pi) = \bar{u}_m(\pi)$. Conversely, if $v(\pi, \delta) = v_0(\pi) = \bar{u}_m(\pi)$ then $\pi \in J_m(\delta)$ and a_m is the optimal choice. As $\bar{u}_m(\pi)$ is an affine function of π , and $v(\cdot, \delta)$ is weakly convex, $J_m(\delta)$ must be an interval.

For the second half, a_1 is myopically strictly optimal for the focused belief $\pi_n = 0$, and since it updates to $\pi_{n+1} = \pi$ a.s. no matter which rule is applied, it is also dynamically optimal for any discount factor $\delta \in [0, 1)$. A similar argument holds when $\pi_n = 1$. \square

Define the expected utility frontier function v_0 by $v_0(\pi) = \max_m \bar{u}_m(\pi)$, that is, the expected utility an individual would obtain by choosing the myopically optimal action.⁹

Claim A.2 (Iterates) *The sequence $\{T_\delta^n v_0\}$ consists of weakly convex functions that are pointwise increasing, and converge to $v(\cdot, \delta)$. The value $v(\cdot, \delta)$ weakly exceeds v_0 .*

Proof: To maximize $\sum_{a_m \in A} \psi(a_m | \pi, x) [(1 - \delta) \bar{u}_m(q(\pi, x, a_m)) + \delta v_0(q(\pi, x, a_m))]$ over x for given π , one possible policy \hat{x} ensures that the myopically optimal signal a_m occurs with probability one. Then $q(\pi, \hat{x}, \hat{x}(\sigma)) = \pi$ a.s., resulting in value $v_0(\pi)$. Optimizing over all $x \in X$, we see that $T_\delta v_0(\pi) \geq v_0(\pi)$ for all π . By induction, $T^n v_0 \geq T^{n-1} v_0$, yielding (as per usual) a pointwise increasing sequence converging to the fixed point $v(\cdot, \delta) \geq v_0$. \square

The following either is or ought to be a folk result for optimal experimentation, but we have not found a published or cited proof of it.¹⁰

Claim A.3 (Monotonicity) *The value function weakly rises when δ increases: Namely, for $\delta_1 > \delta_2$, $v(\pi, \delta_1) \geq v(\pi, \delta_2)$ for all π .*

Proof: Clearly, $\sum_{a_m \in A} \psi(a_m | \pi, x) \bar{u}_m(q(\pi, x, a_m)) \leq \sum_{a_m \in A} \psi(a_m | \pi, x) v(q(\pi, x, a_m))$ for any x and any function $v \geq v_0$. If $\delta_1 > \delta_2$, then $T_{\delta_1} v_0 \geq T_{\delta_2} v_0$, since there is more weight on the larger component of (A-1). Inductively, we have $T_{\delta_1}^n v_0 \geq T_{\delta_2}^n v_0$, since one possible policy under δ_1 is to choose the x optimal under δ_2 . Let $n \rightarrow \infty$ and apply Claim A.2. \square

⁹Observe how this differs from $v(\pi, 0) \equiv \sup_x \sum_m \psi(a_m | \pi, x) \bar{u}_m(q(\pi, a, a_m))$. In other words, $v(\pi, 0)$ allows the myopic individual to observe one signal σ before obtaining the ex post value $v_0(\rho(\pi, \sigma))$.

¹⁰But ABHJ do assert without proof (p. 625) that the patient value function exceeds the myopic one.

Claim A.4 (Inclusion) *All basins weakly shrink when δ increases: In other words, $\forall a_m \in A$, if $1 > \delta_1, \delta_2 \geq 0$, then $J_m(\delta_1) \subseteq J_m(\delta_2)$.*

Proof: As seen in Claims A.3 and A.2, $v(\pi, \delta_1) \geq v(\pi, \delta_2) \geq v_0(\pi) \geq \bar{u}_m(\pi)$ for all π , when $\delta_1 > \delta_2$. For $\pi \in J_m(\delta_1)$, we know $v(\pi, \delta_1) = \bar{u}_m(\pi)$ and thus $v(\pi, \delta_2) = \bar{u}_m(\pi)$. The optimal value can thus be obtained by inducing a_m a.s., so that $\pi \in J_m(\delta_2)$. \square

Claim A.5 (Unbounded Beliefs) *With unbounded private beliefs, only basins for the extreme actions are empty, with $J_1(\delta) = \{0\}$ and $J_M(\delta) = \{1\}$; all other $J_m(\delta)$ are empty.*

Proof: SS establish for the myopic model that all $J_m(0)$ are empty, except for $J_1(0) = \{0\}$ and $J_M(0) = \{1\}$. Now apply Claims A.1 and A.4. \square

Claim A.6 (Bounded Beliefs) *If the private beliefs are bounded, then $J_1(\delta) = [0, \underline{\pi}(\delta)]$ and $J_M(\delta) = [\bar{\pi}(\delta), 1]$, where $0 < \underline{\pi}(\delta) < \bar{\pi}(\delta) < 1$.*

Proof: We prove that for sufficiently low beliefs it is optimal to choose a rule x such that a_1 occurs with probability one; the argument for large beliefs is very similar. Since action a_1 is optimal at belief $\pi = 0$, and is not weakly dominated, there must be some positive length interval $[0, \tilde{\pi}]$ on which $\bar{u}_1(\pi) = v_0(\pi)$, i.e. a_1 is the optimal choice for beliefs $\pi \leq \tilde{\pi}$. Moreover, since each \bar{u}_m is affine, $\exists \underline{u} > 0$ such that $\bar{u}_1(\pi) > \bar{u}_m(\pi) + \underline{u}$ for all $m \neq 1$, and for all beliefs π in the interval $[0, \tilde{\pi}/2]$.

No observation $a \in A$ can ever yield a stronger signal than any $\sigma \in \text{supp}(\mu) \subseteq [\underline{\sigma}, \bar{\sigma}] \subset (0, 1)$. So any initial belief π is updated to at most $\bar{q}(\pi) = \pi\bar{\sigma}/[\pi\bar{\sigma} + (1 - \pi)(1 - \bar{\sigma})]$. For π sufficiently small, $\bar{q}(\pi) \in [0, \tilde{\pi}/2]$ and $\bar{q}(\pi) - \pi$ is arbitrarily small. By the continuity of v , $v(\bar{q}(\pi), \delta) - v(\pi, \delta)$ is then arbitrarily small — in particular, less than $\underline{u}(1 - \delta)/\delta$ for small enough π . The Bellman equation $T_\delta(v) = v$ corresponding to (A-1) reveals that it is strictly suboptimal to risk any outcome other than a_1 for such small beliefs. \square

Claim A.7 (Limiting Patience) *For large enough δ , all basins disappear except for $J_1(\delta)$ and $J_M(\delta)$, while $\lim_{\delta \rightarrow 1} J_1(\delta) = \{0\}$ and $\lim_{\delta \rightarrow 1} J_M(\delta) = \{1\}$.*

Proof: Fix $\delta \in [0, 1)$, and an action index m ($1 < m < M$) for which $J_m(\delta) = [\pi_1, \pi_2]$, for some $0 < \pi_1 \leq \pi_2 < 1$. Since there are informative private beliefs, $\exists x^* \in (1/2, 1)$ with $1 > \mu^H([x^*, 1]) > \mu^L([x^*, 1]) > 0$. We shall consider the alternative policy x , with interval boundaries $\bar{x}_{m-1} = 0$, $\bar{x}_m = x^*$, $\bar{x}_{m+1} = 1$ (see Lemma A.1).

Updating π with observation of event $\{\sigma \in [x^*, 1]\}$ yields the posterior belief $q(\pi) = \pi\mu^H([x^*, 1])/[\pi\mu^H([x^*, 1]) + (1 - \pi)\mu^L([x^*, 1])]$ in state H . For any closed subinterval $I \subset (0, 1)$, in particular one with $I \supseteq J_m(\delta)$, there exists $\varepsilon \equiv \varepsilon(I) > 0$ with $q(\pi) - \pi \geq \varepsilon$ for all $\pi \in I$. By definition of ε , q maps the interval $[\pi_2 - \varepsilon/2, \pi_2]$ into (but not necessarily

onto) $[\pi_2 + \varepsilon/2, 1]$. Choose $\bar{u} > 0$ so large that $\bar{u}_m(\pi) < \bar{u}_{m+1}(\pi) + \bar{u}$ for all $\pi \in [0, 1]$. Since $v(\pi, \delta) > \bar{u}_m(\pi)$ outside $J_m(\delta) = [\pi_1, \pi_2]$, and both are continuous in π , we may also choose $\underline{u} > 0$ so small that $v(\pi, \delta) > \bar{u}_m(\pi) + \underline{u}$ for all $\pi \in [\pi_2 + \varepsilon/2, 1]$. By Claim A.3, we thus have $v(\pi, \delta') > \bar{u}_m(\pi) + \underline{u}$ for all $\delta' > \delta$. If $\delta' > \delta$ is so large that $(1 - \delta')\bar{u} < \delta'\underline{u}$, then the Bellman equation $T_\delta(v) = v$ reveals that our suggested policy x beats inducing a_m a.s. when $\pi \in [\pi_2 - \varepsilon/2, \pi_2]$. By iterating this procedure a finite number of times, each time excising length $\varepsilon/2$ from interval $J_m(\delta)$, we see that $J_m(\delta)$ evaporates for large enough δ .

If $m = 1$ or $m = M$, this procedure can still be applied repeatedly, to show that $J_m(\delta) \cap I$ vanishes for large enough δ for any closed $I \subset (0, 1)$. \square

A.3 Proof of Proposition 3

In light of the interval structure found in Lemma A.1, \mathcal{SP} 's problem is simply to determine the chances $\psi(a_m|\pi)$ with which each action should be chosen (or equivalently, what fraction of the signal space maps into each action). Thus, the choice set is WLOG the compact M -simplex $\Delta(A)$ (that is, the same strategy space as in the original observational learning model). The objective function in the Bellman equation corresponding to (A-1) is continuous in this choice vector and in π , and it follows from the Theorem of the Maximum (e.g. Theorem I.B.3 of Hildenbrand (1974)) that the non-empty correspondence of optimal rules is upper hemi-continuous in π .¹¹

Lemma A.2 (Overturning Principle) *For all $\delta \in [0, 1)$ and $a_m \in A$ with $J_m(\delta) \neq \emptyset$, there exists $\varepsilon > 0$ and an open interval $K \supset J_m(\delta)$, such that $\forall \pi \in K$ and $\forall a \neq a_m$, $\|q(\pi, x, a) - \pi\| > \varepsilon$ when x is the optimal rule.*

Proof: For $\pi \in J_m(\delta)$ we know that the only optimal rule is to a.s. induce a_m . As the correspondence of optimal rules is u.h.c. in π , the optimal rule for π near $J_m(\delta)$ induces actions other than a_m for only the most extreme portion of $\text{supp}(\mu)$. With bounded beliefs, π is bounded away from 0 and 1, and so it is immediate that the public belief must jump at least some $\varepsilon > 0$ upon observing any action other than a_m .

With unbounded beliefs, an extra step is needed. Assume WLOG that π is near 0. Then $q(\pi, x, a_1) \leq \pi$, i.e. the updated belief is close to the current, so that $v(q(\pi, x, a_1), \delta) - v(\pi, \delta)$ is near zero. Choosing an action other than a_1 implies a boundedly positive immediate loss, which must be compensated by gained future value (again, by the Bellman equation). Since v is continuous, it follows that $q(\pi, x, a_m)$ must be boundedly higher than π . \square

¹¹The optimal rule is a point-valued function except when there is an atom on the interval boundary.



References

AGHION, P., P. BOLTON, C. HARRIS, and B. JULLIEN (1991): "Optimal Learning by Experimentation," *Review of Economic Studies*, 58(196), 621-654.

BANERJEE, A. V. (1992): "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 107, 797-817.

BIKHCHANDANI, S., D. HIRSHLEIFER, and I. WELCH (1992): "A Theory of Fads, Fashion, Custom, and Cultural Change as Information Cascades," *Journal of Political Economy*, 100, 992-1026.

EASLEY, D., and N. KIEFER (1988): "Controlling a Stochastic Process with Unknown Parameters," *Econometrica*, 56, 1045-1064.

HILDENBRAND, W. (1974): *Core and Equilibria of a Large Economy*. Princeton University Press, Princeton.

LEE, I. H. (1993): "On the Convergence of Informational Cascades," *Journal of Economic Theory*, 61, 395-411.

MCLENNAN, A. (1984): "Price Dispersion and Incomplete Learning in the Long Run," *Journal of Economic Dynamics and Control*, 7, 331-347.

ROTHSCHILD, M. (1974): "A Two-Armed Bandit Theory of Market Pricing," *Journal of Economic Theory*, 9, 185-202.

RUDIN, W. (1987): *Real and Complex Analysis*. McGraw-Hill, New York, 3rd edn.

SMITH, L., and P. SØRENSEN (1996a): "Pathological Outcomes of Observational Learning," MIT Working Paper.

——— (1996b): "Rational Social Learning Through Random Sampling," MIT mimeo.

STOKEY, N. L., and R. E. LUCAS (1989): *Recursive Methods in Economic Dynamics*. Harvard University Press, Cambridge, Mass.

Date Due

Lib-26-67

